JOURNAL OF BACTERIOLOGY, Jan. 2010, p. 179–190 Vol. 192, No. 1
0021-9193/10/$12.00 doi:10.1128/JB.01174-09


# Genomic and Biological Analysis of Phage Xfas53 and Related Prophages of *Xylella fastidiosa*▽

Elizabeth J. Summer,[1] Christopher J. Enderle,[2] Stephen J. Ahern,[2] Jason J. Gill,[1] Cruz P. Torres,[2] David N. Appel,[2] Mark C. Black,[3] Ry Young,[1] and Carlos F. Gonzalez[2]*

*Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128[1]; Department of Plant Pathology and Microbiology, Texas A&M University, College, Station, Texas 77843-2132[2]; and Texas AgriLife Research and Extension Center, Uvalde, Texas 78801-6205[3]*

Received 31 August 2009/Accepted 27 October 2009

**We report the plaque propagation and genomic analysis of Xfas53, a temperate phage of *Xylella fastidiosa*. Xfas53 was isolated from supernatants of *X. fastidiosa* strain 53 and forms plaques on the sequenced strain Temecula. Xfas53 forms short-tailed virions, morphologically similar to podophage P22. The 36.7-kb genome is predicted to encode 45 proteins. The Xfas53 terminase and structural genes are related at a protein and gene order level to P22. The left arm of the Xfas53 genome has over 90% nucleotide identity to multiple prophage elements of the sequenced *X. fastidiosa* strains. This arm encodes proteins involved in DNA metabolism, integration, and lysogenic control. In contrast to Xfas53, each of these prophages encodes head and DNA packaging proteins related to the siphophage lambda and tail morphogenesis proteins related to those of myophage P2. Therefore, it appears that Xfas53 was formed by recombination between a widespread family of *X. fastidiosa* P2-related prophage elements and a podophage distantly related to phage P22. The lysis cassette of Xfas53 is predicted to encode a pinholin, a signal anchor and release (SAR) endolysin, and Rz and Rz1 equivalents. The holin gene encodes a pinholin and appears to be subject to an unprecedented degree of negative regulation at both the level of expression, with rho-independent transcriptional termination and RNA structure-dependent translational repression, and the level of holin function, with two upstream translational starts predicted to encode antiholin products. A notable feature of Xfas53 and related prophages is the presence of 220- to 390-nucleotide degenerate tandem direct repeats encoding putative DNA binding proteins. Additionally, each phage encodes at least two BroN domain-containing proteins possibly involved in lysogenic control. Xfas53 exhibits unusually slow adsorption kinetics, possibly an adaptation to the confined niche of its slow-growing host.**

The gammaproteobacterium *Xylella fastidiosa* is an insect-transmitted, xylem-inhabiting bacterium and the causal agent of several plant diseases, most notably Pierce's disease of grapes (PD) and citrus variegated chlorosis (CVC) (12). Because of the economic importance of this pathogen, it was the subject of early genomic analysis; the genome of the CVC strain 9a5c was the first completed for a plant pathogen (53). Five additional *X. fastidiosa* genome assemblies are available: a PD strain (Temecula), an oleander leaf scorch (OLS) strain (Ann1), and three almond leaf scorch (ALS)-associated strains (M12, M23, and Dixon) (3, 4, 14, 53). One conclusion from these analyses is that phage may have played a significant role in the evolution of *X. fastidiosa* (19, 44, 53). Phage-related sequences make up between 9% and 15% of the *Xylella* genomes, and 47 phage elements were identified in four genomes (18). Many strain-specific differences are located in prophage (18). Most of the prophage elements appeared to be the result of independent insertion events, implying that active phages are frequently encountered. However, most insertions occur in a common 900-kb region and were associated with large genomic rearrangements (18). Whether the prophages are directly responsible for rearrangements or whether sites prone to recombination are also prone to phage integration is not known.

Despite the abundant evidence of their presence, there are no reports of successful plaque propagation of *X. fastidiosa* phage. Presumptive podophage particles were observed by electron microscopy in ultrathin sections of strain Temecula (13). The majority of prophage sequences were reported to be structural genes for siphophage assembly, and putative siphophage were also observed in an unidentified strain (18). Like many bacteria, the polylysogenic status of *X. fastidiosa* implies that each strain may produce more than one type of phage. Therefore, it is not possible to conclusively correlate an image of a virion produced by a lysogen with a specific prophage region based on morphological descriptions alone, nor is it possible to draw any conclusion about whether any particular observed virion is actually viable. Phages that can be propagated and plaque purified on *X. fastidiosa* are needed to further our understanding of *X. fastidiosa* phage biology. The work presented here describes the propagation, purification, and genomic analysis of *X. fastidiosa* phage Xfas53.

## MATERIALS AND METHODS

**Media, culture conditions, and strains.** PW broth medium as modified by Sherald et al. (52), designated PW-M, was used for growth of *Xylella* isolates, except that the final bovine serum albumin content was 0.3% as modified by Hill and Purcell (27). For solid medium (PW-MG) and soft gel, the PW-M broth was

* Corresponding author. Mailing address: Department of Plant Pathology & Microbiology, Texas A&M University, 2132 TAMU, College Station, TX 77843. Phone: (979) 845-8462. Fax: (979) 845-6483. E-mail: cf-gonzalez@tamu.edu.

▽ Published ahead of print on 6 November 2009.

amended with 9 g/liter and 4.5 g/liter, respectively, of Gelrite Gellan gum (Sigma). Alternative solidifying agents tested included Gelzan CM (Sigma), Phytagel (Sigma), and plant cell culture-tested (PCCT) agar (Sigma). Cultures were grown at 28°C, and liquid cultures were monitored at $\lambda = 600$ nm using sidearm flasks. California *X. fastidiosa* isolates included in the study were Temecula, Ann1, and Dixon (25). Texas isolates included one each from *Plantanus occidentalis* (XF1), *Iva annua* (XF18), *Helianthus annuus* (XF5), *Ratibida columnifera* (XF37), *Ambrosia psilostachya* (XF23), *Vitis aestivalis* (XF39), and *Vitis mustangensis* (XF41); three isolates from *Ambrosia trifida* var. *texana* (XF16, -40, and -43); two from *Nerium oleander* (XF93 and -95); and 15 from *Vitis vinifera* (XF48, -50, -52, -53, -54, -56, -58, -59, -60, -66, -67, -70, -71, -76, and -78). All isolates were single colony purified by the streak isolation method and stored at −80°C after amending PW-M broth cultures to a final concentration of 20% (vol/vol) glycerol. Isolates were confirmed at the species and subspecies level using PCR analysis as previously described (26).

**Bacteriophage production, titration, and single plaque purification.** Cultures were grown in PW-M broth at 28°C for 7 to 10 days. Cells were removed by centrifugation (10,000 × *g*, 15 min at 5°C) in a J2-21 centrifuge (Beckman Coulter). The supernatant was filtered through a 0.22-μm filter (Supor; Pall Life Sciences). Five- to seven-day cultures of *X. fastidiosa* isolates were used to make indicator suspensions in PW-M broth ($A_{600} = 0.5$). Soft gel overlays used to survey bacterial supernatants for phage activity were made by mixing 100 μl of the bacterial suspension with 7 ml of molten PW-M soft gel, pouring the mixture on PW-MG plates, and allowing it to solidify and dry. Each supernatant was serially diluted in phage buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 8 mM $MgSO_4$), and soft gel overlays were spotted with 10 μl of each dilution, allowed to dry, and incubated at 28°C for 5 to 7 days. Positive supernatants showing either plaques or cleared zones were serially diluted in phage buffer and then plated as described above, except that the phage dilutions were directly mixed with the indicator suspension before addition of molten soft gel. Individual plaques were excised from the overlay and suspended in phage buffer, and their titers were determined. This procedure was repeated twice to obtain a single plaque isolate for each phage. High-titer lysates ($10^{10}$ PFU/ml) were prepared by harvesting overlays of plates exhibiting confluent lysis with 5 ml of phage buffer, macerating the soft gel overlay, clearing the lysate by centrifugation (10,000 × *g*, 15 min at 4°C), and filter sterilizing it through a 0.22-μm filter. Lysates were stored at 4°C.

**CsCl purification.** Filter-sterilized phage suspensions were concentrated by centrifugation (90,000 × *g* for 2.5 h at 5°C) using a type 60Ti rotor in a Beckman L8-70 M ultracentrifuge. Pellets were resuspended in phage buffer and treated with DNase I and RNase A (Sigma) at a final concentration of 1 μg/ml at 25°C for 2 h. CsCl was added to the phage suspension at a final concentration of 0.75 g/ml and centrifuged (300,000 × *g* for 18 h at 5°C) in a VTi 65.2 rotor. The visible phage band was extracted using an 18-gauge syringe needle and dialyzed against phage buffer amended to 1 M NaCl overnight at 4°C and twice for 4 h at 25°C against phage buffer to obtain a suspension of $10^{11}$ PFU/ml. The CsCl-purified phage was stored at 4°C.

**Phage adsorption assay.** The kinetics of phage adsorption was determined by infecting *X. fastidiosa* strain Temecula suspended in PW-M broth at a multiplicity of infection of ~0.01. Samples were taken at 2-h intervals, and titers were determined after removal of cells by filtration of the sample through a 0.2-μm filter using strain Temecula as the indicator. The rate of phage particle disappearance is defined as $dP/dt = -kBP$, where $B$ is the concentration of bacteria, $P$ is the concentration of free phage at any time ($t$), and $k$ is the adsorption constant in ml cell$^{-1}$ min$^{-1}$ (50a).

**Prophage induction.** Three-day cultures ($A_{600} = 0.15$) were washed by centrifugation (10,000 × *g*, 15 min at 5°C), resuspended in 0.85% NaCl, and exposed to 254-nm UV radiation (360 μW/cm$^2$) for 0, 5, 7, and 10 s in a sterile petri dish (5 ml/dish). After exposure, cells were collected by centrifugation, resuspended in PW-M broth, and incubated at 28°C with shaking (200 rpm). One-milliliter samples were removed 1, 2 and 3 days post-UV exposure and filter sterilized, and their titers were determined to determine PFU.

Mitomycin C inductions were performed by adding mitomycin C (Sigma) to a final concentration of 0 to 2.0 μg/ml to growing cultures and incubating them in the dark at 28°C for 48 h with shaking (200 rpm). After 48 h, the titers of filter-sterilized samples were determined using the indicator strain Temecula.

Heat induction experiments were conducted by exposing 3-day cultures grown in PW-M broth at 28°C with shaking (200 rpm) ($A_{600} = 0.15$) to 42°C for 0, 15, 30, and 60 min in a shaking water bath. Cultures were then incubated at 28°C with shaking (200 rpm) for an additional 4 days and filter sterilized, and their titers were determined.

**TEM.** Electron microscopy of CsCl-purified phage ($3.7 \times 10^{11}$ PFU/ml) was performed by diluting the phage with phage buffer and applying 5 μl onto a freshly glow-discharged Formvar-carbon-coated grid for 1 min. Grids were then washed briefly on deionized water drops and stained with 2% (wt/vol) aqueous uranyl acetate. For transmission electron microscopy (TEM) of phage with bacteria, a bacterial suspension ($8 \times 10^8$ CFU/ml) was mixed with dilutions of purified phage and incubated at 25°C for 30 min. Five microliters of the mixtures was then applied onto freshly glow-discharged Formvar-carbon-coated grids and left to adhere for 1 min. Grids were then washed briefly on deionized water drops and stained with 2% (wt/vol) aqueous sodium phosphotungstate, pH 7.0. Specimens were observed on a JEOL 1200EX transmission electron microscope operating at an acceleration voltage of 100 kV. Images were recorded at calibrated magnifications on Kodak 4489 film.

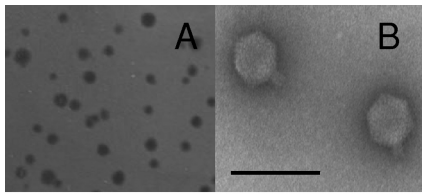


FIG. 1. (A) Plaques formed by phage Xfas53 on a soft gel (Gelrite, Gellan gum) overlay of *X. fastidiosa* strain Temecula. (B) Electron micrograph of CsCl-purified phage Xfas53 negatively stained with 2% (wt/vol) aqueous uranyl acetate. Bar = 100 nm.

**Phage DNA isolation and genome sequencing.** DNA was isolated from CsCl-purified phage suspensions as described previously (54). Genomic DNA was sized by pulsed-field gel electrophoresis. Library preparation, sequencing, sequence assembly, and analyses were performed essentially as described previously (54, 56, 57). The genomic DNA library was prepared in the pSmart-LCKan vector (Lucigen). Sequencher (Gene Codes Corporation) was used for sequence assembly and editing. Protein coding regions were predicted using Genemark (http://opal.biology.gatech.edu/GeneMark/gmhmm2_prok.cgi) and manually edited in Artemis (http://www.sanger.ac.uk/Software/Artemis/) using the phage genome annotation tool ArtAnnoPipe (http://athena.bioc.uvic.ca/node/541) (42, 50). Dot plots were generated using JDotter (6). Predicted proteins were compared to proteins in the GenBank database using BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Conserved domains, lipoprotein processing signals, and transmembrane domains (TMDs) were identified with InterProScan (http://www.ebi.ac.uk/Tools/webservices/services/interproscan), LipoP (http://www.cbs.dtu.dk/services/LipoP/), and TMHMM (http://www.cbs.dtu.dk/services/TMHMM/), respectively.

**Protein analysis.** SDS-PAGE was carried out on CsCl-purified phage particles along with Precision Plus protein standard (Bio-Rad) and stained with Coomassie brilliant blue R250 (37).

**Nucleotide sequence accession number.** The sequence of phage Xfas53 has been entered as GenBank accession number GQ421471.

## RESULTS

**Phage plating, isolation, and morphology.** The growth of *X. fastidiosa* lawns and phage plaque formation were found to be sensitive to the solidifying agent used in the medium. *X. fastidiosa* failed to form uniform lawns in standard Bacto Agar. Alternative solidifying agents were tested, including Gelrite, Gelzan, and PCCT agar. Plaque formation was observed in overlays made with Gelrite at 0.45% (Fig. 1A) or PCCT agar at 0.75%. Very small plaques were observed with Gelzan and Phytagel (data not shown). Experiments described here were performed with Gelrite (Gellan gum) overlays. Supernatants from liquid cultures of 30 individual *X. fastidiosa* isolates were spotted on overlays seeded with each isolate. Supernatants from four of these (isolates 48, 53, 58, and 60) developed plaques on overlays of strain Temecula following a 5- to 6-day incubation. Titers in these uninduced culture supernatants ranged from $1 \times 10^3$ to $5 \times 10^3$ PFU/ml.

The phage produced by lysogenic *X. fastidiosa* strain 53 was

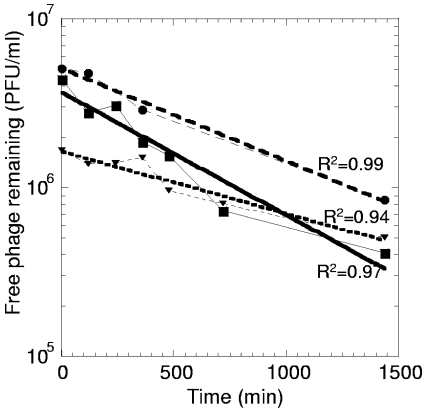


FIG. 2. Adsorption kinetics of phage Xfas53 on *X. fastidiosa* strain Temecula. Data are shown from three independent replicates (indicated by thin lines with squares, circles, and triangles) with regression curves shown as heavy lines, corresponding to data by line type (short-dash, long-dash, or solid lines). Regression curves fit to the data with $R^2$ values ranging from 0.938 to 0.988. Binding rate constants ($k$) were calculated from the slopes of the regression curves, given that the slope is equal to ($-kB$) where $B$ is equal to the concentration of bacteria in CFU/ml.

chosen for further study. Individual plaques were excised and plaque purified twice to obtain a clonal isolate, designated phage vB_XfaP_Xfas53 (Xfas53) (Fig. 1A). After a 5- to 6-day incubation, Xfas53 formed clear plaques on overlays of strain Temecula. An additional 5- to 6-day incubation resulted in the development of turbidity in the plaques, suggestive of lysogen formation; however, viable bacteria could not be rescued from the plates to confirm this (data not shown). Phage Xfas53 failed to form plaques on overlays of strain 53, as would be expected for a superinfection-resistant lysogen. Transmission electron microscopy of purified phage revealed that phage Xfas53 possesses a 55-nm head and a short, 12-nm-diameter, noncontractile tail, typical of podophage morphology (Fig. 1B). Electron micrographs of concentrated supernatants revealed that the phage in the other three positive supernatants also possessed podophage morphology (data not shown). The adsorption rate constant of phage Xfas53 for host *X. fastidiosa* Temecula was determined from three replicate experiments to be $(6.8 \pm 1.5) \times 10^{-12}$ ml cell$^{-1}$ min$^{-1}$ (Fig. 2).

Prophage induction experiments of *X. fastidiosa* strain 53 using either mitomycin C, UV, or heat shock treatment did not result in an increase in phage production above that observed in uninduced cultures as determined by phage titer or culture density. Mitomycin C treatments above 1.5 μg/ml resulted in a complete loss of cell viability without inducing the production of virions, as indicated by the absence of changes in culture turbidity following treatment. In heat induction experiments, there was no increase in the titer above that observed in the control culture. The exposure of cells to 42°C for 30 and 60 min resulted in loss of cell viability, as measured by absorbance. Examination by electron microscopy of a 100-fold-concentrated supernatant of uninduced strain Temecula, grown under the same conditions in which strain 53 produced phage Xfas53, showed no evidence of phage particles.

**Overview of the Xfas53 genome.** Xfas53 genomic DNA migrated as a single 36-kb band in pulsed-field gel electrophoresis (data not shown). Clones from a shotgun library were sequenced to eightfold coverage, and the sequence reads were assembled into a single contig of 36,674 bp. The genome has an average GC content of 57%, which is slightly higher than the 53% GC content of *Xylella*. The circular assembly was opened upstream of the predicted integrase (designated *gene01*). The longest region not predicted to encode a protein is 820 nucleotides between *gene23* and *gene24*. A GC skew plot (not shown) indicated that this gap either corresponds to a replication origin or is possibly a relic of the recombinatorial origin of the right and left arms of the phage (see below).

A total of 45 protein coding genes were predicted to be encoded by the Xfas53 genome (Fig. 3; Table 1). Xfas53 genes are organized into two transcription units with the first 18 genes transcribed from the reverse strand and all but three of the remaining 27 genes transcribed from the forward strand (Fig. 3). Genes at the presumptive divergent promoter region included some with similarity to proteins implicated in lysogenic control.

**Functional assignments for predicted Xfas53-encoded proteins. (i) Integration and lysogenic control.** Xfas53 is a temperate phage, and among the predicted genes are those expected to be responsible for phage integration and regulation of lysogeny. *gene01* is predicted to encode an integrase, as gp01 has a conserved catalytic C-terminal integrase domain and sequence similarity to the archetypical integrase from phage lambda. There are four or five Xfas53 gp1 homologues exhibiting over 90% identity in each of the *X. fastidiosa* genomic assemblies. *Xylella* strains are known to possess multiple, closely related integrases (18). For example, strain Temecula possesses five integrases (locus tags PD1019, PD0764, PD1196, PD0384, and PD1139) that are over 90% identical to Xfas53 gp1. These, along with the Xfas53 integrase, cluster with the largest *Xylella* integrase groups (18). The most closely related phage-encoded protein is only 52% similar, encoded by *Pseudomonas* phage F10 (36).

gp17 was identified as a lambda cI-like repressor based on the presence of a conserved C-terminal LexA-family serine peptidase domain (PFAM PF00717 peptidase_S24). This signature belongs to proteins of the MEROPS peptidase family, which includes numerous prophage repressor proteins, including *Salmonella* (SE1 and ST104) and *Pseudomonas* (F10, MP22, DMS3, and D3) phages. The D3 cI homologue was shown to exhibit binding specificity for the D3 operator sequences (21, 34). Xfas53 also encodes two BroN domain-containing proteins, gp08 and gp22 (29). BroN domains are DNA-interacting domains found in proteins whose specific function is conferred by a number of different C-terminal domains. A third protein, gp21, possesses a COG3617 antirepressor domain. The significance of these proteins is discussed in more detail below.

**(ii) DNA replication.** *gene04*, *gene06*, and *gene07* encode proteins with conserved domains indicative of a DNA helicase, nuclease, and DNA polymerase, respectively. Highly similar homologues of these three genes, as well as *gene09* and *gene10*, are present in syntenic clusters in several otherwise unrelated podophages. These include *Burkholderia* phage BcepC6B, *Acyrthosiphon pisum* endosymbiont phage APSE-1, and *Bordetella* phages BPP-1, BMP-1, and BIP-1. This linkage suggests that gp09 and gp10 also function in DNA replication. *gene23*

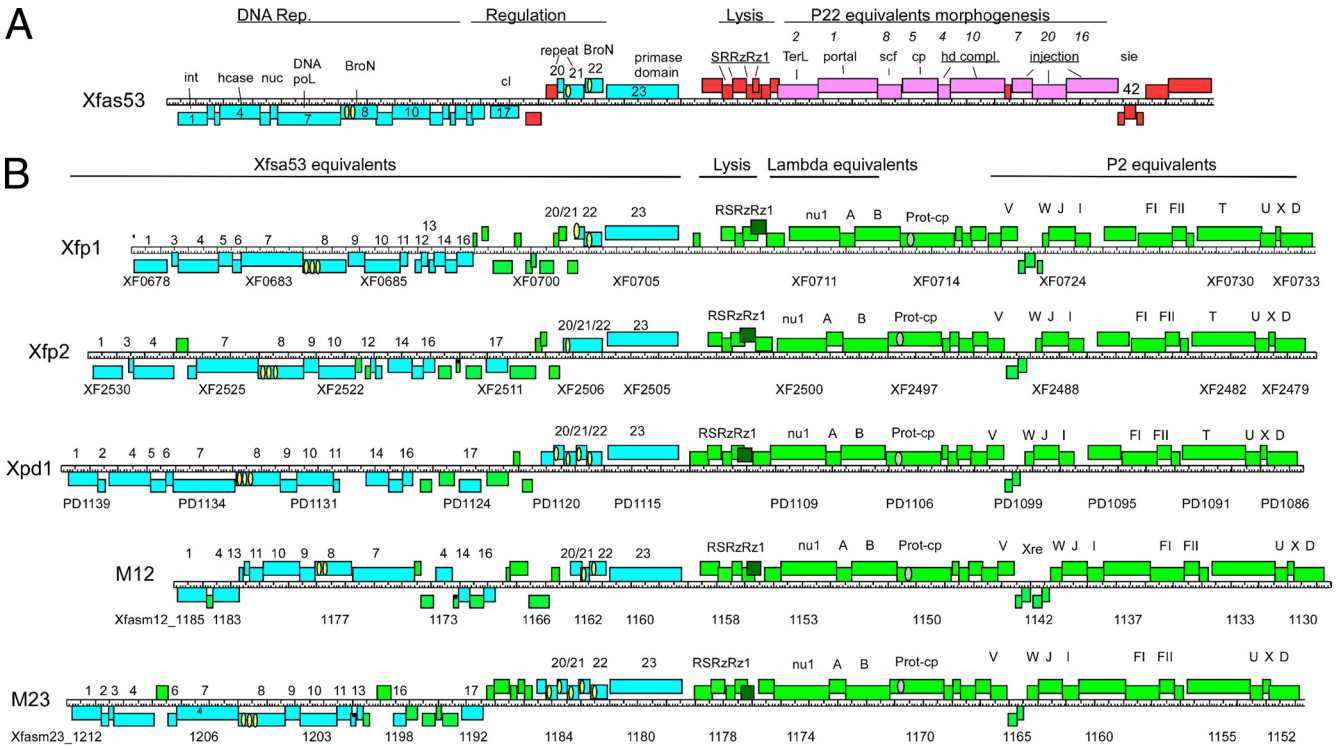


FIG. 3. Genome maps of Xfas53 and Xfp1-like prophage elements. (A) Genome map of phage Xfas53 with genes drawn to scale on forward and reverse strands. Homologues of genes indicated in blue are also found in the Xfp1-like prophage elements. Homologues of genes indicated in pink are present in phage P22. Genes indicated in red do not have homologues in P22 or Xfp1. Yellow ovals indicate BroN and/or COG3617 domains. Abbreviations: int, integrase; hcase, helicase; nuc, nuclease; DNA pol, DNA polymerase; TerL, terminase large subunit; scf, scaffold; cp, capsid protein; hd compl, head completion proteins; injection, DNA injection protein; sie, superinfection exclusion. (B) Maps of prophage elements Xfp1, Xfp2, Xpd1, M12, and M23. Locus tags are indicated for select genes. Genes shown in blue have a homologue present in Xfas53 (Xfas53 gene number is indicated above). Genes indicated in green do not have a homologue in Xfas53. Genes for lambda head assembly and P2 tail assembly are indicated. Yellow ovals indicate BroN and/or COG3617 domains. The gray oval in the Prot-cp indicates capsid protein with protease domain. Dark green indicates corrected Rz1 coding regions.

encodes an 845-amino-acid (aa) residue protein with an amino terminus that aligns only with highly related prophage-encoded proteins in the *Xylella* genome entries and BCPG_00731 from *Burkholderia cenocepacia* PC184 (data not shown). The carboxy-terminal half of gp23 contains both a D5_N domain and a P4 primase C-terminal domain. These latter two domains are frequently associated with N-terminal primase domains (30). The lack of a detectable primase active site conserved domain at the amino terminus of gp23, however, indicates that this protein cannot be conclusively annotated as a primase without experimental evidence.

**(iii) Host cell lysis.** The Xfas53 lysis cassette consists of *gene25*, *gene26*, *gene27*, and *gene28* and is organized similarly to the archetypical lambda lysis cassette *SRRzRz1*. Of these, only *gene26*, encoding the endolysin (Lyz), was identified based on amino acid sequence similarity. The Xfas53 Lyz is related to members of the canonical T4 lysozyme family, with a glycoside hydrolase domain containing an E-8 aa-D/C-5 aa-T catalytic triad that flanks a substrate-binding cleft (Fig. 4A). However, Xfas53 Lyz has a predicted TMD at the N terminus (residues 12 to 30), indicating that it belongs to the signal anchor and release (SAR) endolysin subgroup (65). SAR endolysins are secreted by the host *sec* system to the periplasm, where they accumulate during the latent period in an inactive form tethered to the membrane by the SAR domain. Like other SAR domains and unlike normal TMDs, the SAR sequence of Xfas53 Lyz is only weakly hydrophobic, with only 10 of its 19 positions occupied by strongly hydrophobic amino acid residues, a disproportion that is required for the ability to exit the bilayer (58). When the membrane is permeabilized (by the holin) and thus deenergized, the SAR domain exits the bilayer, refolds to its active conformation, and attacks the peptidoglycan (Fig. 4B). In the membrane-tethered form, Glu31 of the catalytic triad of Xfas53 Lyz is predicted to be juxtaposed to the periplasmic surface of the cytoplasmic membrane, making it impossible to form a functional active site (Fig. 4A). Xfas53 Lyz thus belongs to a class of SAR endolysins which rely on the steric hindrance of the membrane to prevent premature activation of the endolysin, rather than the well-documented covalent control based on disulfide bond isomerization observed in the SAR endolysin of coliphage P1 (64).

Xfas53 *gene25*, or *hol*, is predicted to encode the holin. In addition to its position preceding the endolysin gene, *hol* has other hallmarks of a holin gene (62). First, Xfas53 Hol is the right size, spanning 115 residues, and has an extremely hydrophilic, cytoplasmic C-terminal domain. Second, the Xfas53 Hol possesses two TMDs, the most hydrophobic of which is TMD2 (residues 73 to 92) (Fig. 4C). Another TMD, not predicted by

TABLE 1. Xfas53-encoded gene products[a]

| GP | Annotation | MW | Representative non-*Xylella* BLAST hit [gene product; organism; E value] (membrane localization signals) | Xpd1, Xop2 homolog locus tags | Conserved domains |
|---|---|---|---|---|---|
| gp01 | Integrase | 38,356 | [Integrase; phage F10; 4e-48] | PD1139, XOP_1071 | PF00589 Phage_integrase, cd00800, INT |
| gp02 | Xf phage prt | 9,472 | None | PD1138 | None |
| gp03 | Xf phage prt | 6,037 | None | None | None |
| gp04 | Helicase, Bbp47 hom | 53,310 | [Bbp47; phage BPP-1; 0] [helicase; phage BcepC6B; 3e-140] | PD1137, XOP_1069 | (N-terminal cl09099, P-loop NTPase, PF00176 SNF2_N ) (C-terminal cd00079, helicase) |
| gp05 | Xf phage prt | 12,666 | None | PD1136 | None |
| gp06 | nuclease, Bbp43 hom | 10,308 | [Bbp43; phage BPP-1; 9e-28] [gp30; phage BcepC6B; 4e-14] | PD1135, XOP_1066 | PF08774 VRR_NUC restriction endonuclease-like fold families |
| gp07 | DNA pol | 79,317 | [Bbp42 phage BPP-1; 0] [DNA pol; phage BcepC6B; 5e-178] | PD1134, XOP_1064 | (N-terminal cl10012, DnaQ-like exonuclease) (C-terminal cl02626, DNA_pol_A) |
| gp08 | BroN domain prt | 47,564 | [MPMin1_gp10; phage Min1; 2e-04] | PD1133, XOP_1065 | (pfam02498, Bro-N, COG3617, prophage antirepressor) two domains |
| gp09 | Bbp40 hom | 20,346 | [Bbp40; phage BPP-1; 9e-72] [gp38; phage BcepC6B; 4e-20] | PD1132 | PRODOM PD339480; PHAGE. PHAGE-TYPE |
| gp10 | Bbp38 hom | 46,805 | [Bbp38; phage BPP-1; 1e-129] [gp39; phage BcepC6B; 5e-72] | PD1131, XOP_1063 | (Weak relationship to TIGR00372, cas4, CRISPR-associated prt Cas4) |
| gp11 | Xf phage prt | 17,249 | None | PD1130, XOP_1062 | None |
| gp12 | Xf phage prt | 7,249 | None | None | None |
| gp13 | Xf phage prt | 7,812 | None | None | None |
| gp14 | Xf phage prt | 15,414 | None | PD1129, XOP_1061 | None |
| gp15 | Xf phage prt | 6,992 | None | None | None |
| gp16 | Xf phage prt | 15,538 | None | PD1127, XOP_1060 | None |
| gp17 | cI repressor | 37,238 | [Repressor; phage F10; 2e-10] | PD1124, XOP_1059 | PF00717; Peptidase_S24, cd06529, LexA |
| gp18 | Phage prt | 18,746 | [DMS3-25; phage DMS3; 0.002] | PD1122 | None |
| gp19 | Xf phage prt | 14,250 | [Hyp prt Rpic_2329; *R. pickettii*; 2e-21] | PD1121 | None |
| gp20 | Xf phage prt | 9,492 | None | PD1120, XOP_1058 | None |
| gp21 | COG3617 antirepressor | 23,300 | [COG3617; prophage antirepressor; *M. magnetotacticum* MS-1; 0.047] | PD1119, XOP_1057 | (COG3617, prophage antirepressor) two domains |
| gp22 | BroN domain prt | 20,949 | [antirepressor; phage F10; 1e-17] | PD1118, XOP_1056 | (pfam02498, Bro-N, COG3617, antirepressor) |
| gp23 | Primase domain prt | 93,821 | [primase; phage A2; 7e-62] | PD1117, XOP_4317 | PF08706 D5_N phage primase_C-terminal, COG3378, phage associated DNA primase |
| gp24 | Xf phage prt | 25,999 | [Hyp prt Smal_2512; *S. maltophilia* R551-3; 5e-47] | XOP_4318 | TIGR02349, DnaJ, COG0484, DnaJ |
| gp25 | Holin | 12,712 | None (TMD 55–70: 73-91) | None | None |
| gp26 | SAR endolysin | 17,823 | [Lysozyme; phage 21; 1e-25] (SAR 13-30) | None | pfam00959, Phage_lysozyme, COG3772, muraminidase, cd00737, endolysin_autolysin |
| gp27 | Rz | 19,285 | [PY54p62; phage PY54; 0.024] (TMD 9-31) | None | None |
| gp28 | Rz1 | 7,774 | None (SPII 13-35) | None | None |
| gp29 | Xf phage prt | 11,416 | None | None | None |
| gp30 | Con hyp prt | 11,415 | [RCOM_0392400; *Ricinus communis*; 1e-11] | None | None |
| gp31 | TerL, P22 gp2 hom | 53,633 | [TerL; phage P22; 3e-105] | XOP_4322 | pfam03237, terminase, COG5565, bacteriophage terminase large (ATPase) subunit |
| gp32 | Portal, P22 gp1 hom | 76,880 | [BA3_0015; phage BA3; 5e-99] [portal; phage HK620; 2e-45] | Xop_4323 | None |

*Continued on following page*

TABLE 1—*Continued*

| GP | Annotation | MW | Representative non-*Xylella* BLAST hit [gene product; organism; E value] (membrane localization signals) | Xpd1, Xop2 locus tag and homologues | Conserved domains |
|---|---|---|---|---|---|
| gp33 | Scaffold, P22 gp8 hom | 31,765 | [BA3_0014; phage BA3; 8e-05] [scaffold; phage P22; 0.046] | Xop_4324 | None |
| gp34 | Major head Prt, P22 gp5 | 43,549 | [BA3_0013; phage BA3; 1e-25] [cp; phage HK620; 0.006] | Xop_4325 | None |
| gp35 | Head comp., P22 gp4 hom | 14,674 | [DNA stabilization; phage HK620; 1e-5] | Xop_4326 | None |
| gp36 | Head comp., P22 gp10 hom | 69,859 | [BA3_0009; phage BA3; 3e-53] [head completion; phage P22; 3e-04] | Xop_4327 | PRODOM PD142451 Q76H18_VVVVV; DNA stabilization packaged late head gene |
| gp37 | Hyp, novel prt | 8,253 | None | None | None |
| gp38 | Injection prt, P22 gp7 hom | 25,349 | [BA3_0006; phage BA3; 5e-09] | Xop_4328 | PRODOM PD104351 VG07_BPST6_Q8HAD9; DNA transfer gp7 injection |
| gp39 | Injection prt, P22 gp20 hom | 42,917 | [gene 12 prt; phage Sf6; 5e-05] | Xop_4329 | None |
| gp40 | Injection prt, P22 gp16 hom | 62,595 | [gene 13 prt; phage Sf6; 5e-12] | Xop_4330 | None |
| gp41 | Hyp novel prt | 9,246 | None (TMD 20-42) | None | None |
| gp42 | Membrane prt | 12,306 | [inner membrane prt; *E. coli* E22; 5e-5] (TMD 61-83, 88-105) | None | COG5346, predicted membrane prt, pfam10097, DUF2335, predicted membrane prt |
| gp43 | Xf phage prt | 9,343 | None | None | None |
| gp44 | Xf phage prt | 27,421 | None | XOP_4331 | None |
| gp45 | Con hyp | 51,607 | [gp87; phage phiJL001; 6e-05] | XOP_4332 | None |

[a] Abbreviations: Xf, *Xylella fastidiosa*; hyp, hypothetical protein; hom, homologue; con, conserved; prt, protein; TMD, transmembrane domain; SPII, signal peptidase II signal sequence; SAR, signal anchor and release domain; *R. pickettii*, *Ralstonia pickettii*; *M. magnetotacticum*, *Magnetospirillum magnetotacticum*; *S. maltophilia, Stenotrophomonas maltophilia*; *E. coli*, *Escherichia coli*.

typical TMD search algorithms, is present from residues 55 to 70 and has characteristics of a SAR domain, with a high percentage (10 out of 16) of weakly hydrophobic or polar residues (65). The presence of a SAR domain followed by a typical TMD suggests that Xfas53 Hol is a pinholin, analogous to the holin of phage 21, $S^{21}$. Pinholins are a recently discovered class of holins that make small holes in the host membrane (47). These holes are sufficient to depolarize the membrane and allow exit of SAR endolysins but are not large enough to allow escape of canonical cytoplasmic endolysins. Even more so than with other holin genes, the structure of *gene25* suggests multiple levels of negative regulation at the levels of transcription, translation, and function. There are three possible start codons in Xfas53 *hol*, all of which have appropriately placed Shine-Dalgarno sequences. Starts at codon 1, 16, and 35 would produce 115 (Hol115)-, 90 (Hol90)-, and 70 (Hol70)-amino-acid products, respectively. The two longer products would have N-terminal extensions with either 1 or 2 extra positively charged residues compared to Hol70. The prototype pinholin, $S^{21}$, exhibited dynamic membrane topology, in that TMD1 is required to exit the bilayer before pinhole formation by TMD2 can proceed. N-terminal extensions with positively charged residues were shown to inhibit escape of TMD1 from the bilayer and not only retarded hole formation but also conferred negative-dominant antiholin character (47). Thus, both Hol115 and Hol90 are predicted to be antiholins that would specifically inhibit Hol70. In addition, the *hol* mRNA would contain a strongly predicted RNA stem-loop structure overlapping the ribosome-binding site for the start codon at 35 (Fig. 4C). This mirrors similar stem-loops that reduce holin translation, in favor of antiholin synthesis, in lambda and other phages (62). Finally, the same stem-loop structure is followed by an oligopyrimidine stretch, strongly suggesting that it constitutes a rho-independent transcriptional terminator.

The spanin component genes, *gene27* (*Rz*) and *gene28* (*Rz1*), were easily identified by their characteristic gene architecture, in which *Rz1* is a reading frame encoding a short (45-aa mature length), proline-rich (14/45 residues) outer membrane (OM) lipoprotein entirely embedded out of frame within *Rz* (55). Like other Rz proteins, Xfas53 Rz is a type II (N-in, C-out) cytoplasmic membrane protein with a highly charged periplasmic domain predicted to have significant α-helical character. The Rz and Rz1 proteins form the spanin complex which spans the periplasm during the latent period (2). When the cell wall is degraded by the endolysin, the complex is thought to undergo a conformational change and cause disruption of the OM, possibly by fusing it with the cytoplasmic membrane.

**(iv) Morphogenesis: relationships of Xfas53 to P22.** The right arm of the Xfas53 genome possesses a virion morphogenesis cassette with protein and gene order similarity to podophage P22, APSE-1, *Shigella flexneri* phage Sf6, and *Thalassomonas* phage BA3 (Fig. 3A) (10, 20, 59, 60). As expected, the Xfas53 virion morphogenesis genes exhibit a mosaic relationship with these phages (10). Individual similarities were not high. For example, the homologues from phage BA3, the most closely related phage, ranged from 44% to 53% in similarity. Because phage P22 has been extensively studied in terms of capsid morphogenesis, DNA packaging, and DNA injection apparatus assembly, it is useful to make comparisons with Xfas53, even when individual gene similarities are low (32). Nine Xfas53-encoded proteins could be confidently correlated with P22 equivalents. These include portal, scaffold,

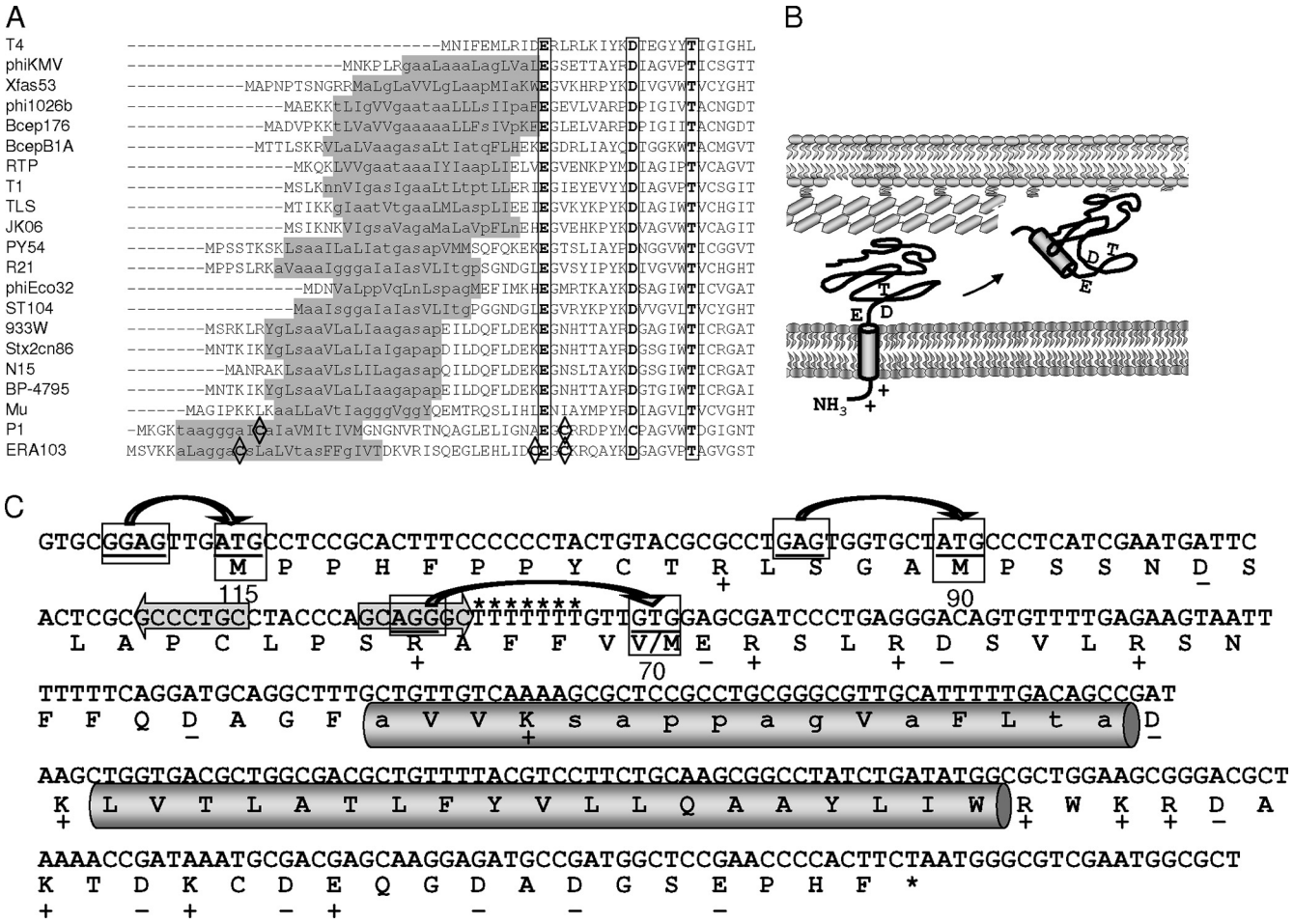


FIG. 4. Holin and endolysin proteins of phage Xfas53. (A) The amino termini of lysozyme homologues from the indicated phages are shown aligned by their conserved catalytic triad (boxed, E-8 aa-D/C-5 aa-T). The SAR domains are indicated in gray. Amino acid residues in the SAR domain that are predicted to be weakly hydrophobic are shown in lowercase. Cysteine residues known to participate in lysozyme activation through disulfide bond isomerization are shown inside diamonds. (B) Model for the activation of Xfas53 lysozyme. When anchored in the inner membrane, the catalytic triad (EDT) is inactive. Release from the membrane causes refolding and activation. (C) The holin gene of Xfas53 suggests multiple gene products. Xfas53 *hol* (*gene25*) is shown with the predicted translation below. Charged residues are indicated by + or −. The two TMDs are indicated with gray cylinders. The three potential translational start codons, indicated at residues 1, 35, and 70, and corresponding Shine-Dalgarno sequences are indicated by boxes and linked by arrows. The potential stem-loop repressor structure is indicated with gray arrows, and the oligopyrimidine tract characteristic of a rho-independent terminator is marked with stars.

major head protein, the three DNA injection proteins, the large terminase subunit (TerL), and two of the three P22 tail accessory factors.

Xfas53 encodes homologues of P22 gp1 (portal), gp8 (prohead scaffold), and gp5 (major capsid); therefore, Xfas53 gp32, gp33, and gp34 probably function as the portal, scaffold, and major head protein, respectively (Fig. 3A; Table 1). The predicted molecular mass of the gp34 major head protein is 43 kDa. When proteins from purified Xfas53 virions were analyzed by SDS-PAGE, a single major band of approximately 43 kDa was observed (data not shown). In P22, three proteins, gp7, gp20, and gp16, are packaged inside the mature capsid and function in DNA injection. Homologues of these proteins were assigned to gp38, gp39, and gp40. Although Xfas53 gp37 is significantly related to P22 gp7, the P22 gp20 and P22 gp16 equivalents are more similar to those from phage Sf6 (10).

DNA packaging is the next step of P22 virion morphogenesis (9, 31). The putative TerL homologue is gp31, which is highly similar (60%) to P22 *gene2*-encoded TerL. Several lines of evidence suggest that Xfas53 utilizes a head-full packaging mechanism, similar to P22 (11), including the lack of defined termini in the sequence assembly, the failure of the genomic DNA to form multimers during agarose gel electrophoresis, and the presence of a *pac* fragment following restriction digest analysis of genomic DNA (not shown). The lack of an identifiable TerS subunit was not surprising, given the diversity of TerS proteins. The size and genetic location of *gene30* near the *pac* site make it a likely *terS* candidate (63).

The most significant differences between Xfas53 and P22 virion morphogenesis were found in the head completion genes (corresponding to P22 gp4, gp10, and gp26) and tailspike gene (corresponding to P22 gp9) (45). Association of P22 gp10 with gp4 is a prerequisite for the attachment of gp26 (45). P22 gp26 forms the central tail needle, which penetrates the host

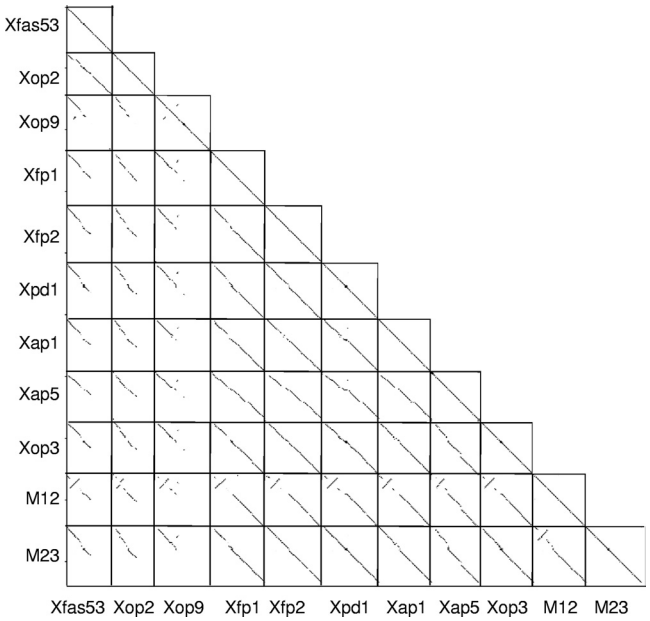


FIG. 5. Dot plot alignments of Xfas53 and *Xylella* prophage elements.

envelope during the DNA injection process, and the gp9 tailspike, which functions in two critical roles: recognition of the host O antigen and hydrolysis of the lipopolysaccharide (45). Xfas53 encodes recognizable homologues of P22 gp4 and gp10 (Xfas53 gp35 and gp36, respectively) but not homologues of either P22 gp26 or gp9. This is not unexpected given that gp9 and gp26 equivalents vary even among the closely related P22-like phages (10). While Xfas53 gp35 is similar in size to P22 gp4, the P22 gp10 equivalent, Xfas53 gp36, is larger (642 residues compared to 472) due primarily to the presence of a carboxy-terminal extension. One possibility is that Xfas53 gp36 is a functional fusion between P22 gp10 and gp26 equivalents. gp36 is predicted to adopt considerable extended conformation (data not shown). Alternatively, due to gene position and size, it is possible that either gp44 or gp45 encodes tailspike or needle equivalents. Structural prediction programs suggest that Xfas53 gp45 might have a tendency to form beta wrap structures, similar to P22 gp26 and gp9 (data not shown).

**(v) Superinfection exclusion.** Embedded within the virion morphogenesis cassette and encoded on the complementary strand are three small open reading frames, *gene41*, *gene42*, and *gene43*, encoding proteins of 84, 116, and 81 residues, respectively (Fig. 3A). Both gp41 and gp42 are strongly predicted to be anchored in the inner membrane by one and two TMDs, respectively (data not shown). gp42 is significantly related to numerous bacterial inner membrane proteins of unknown function. Many phages, including P22, T4, and P1, possess superinfection-exclusion mechanisms that utilize small, poorly conserved, inner membrane or periplasmic proteins to prevent the entry of DNA from superinfecting phage (28, 40, 41, 43, 48, 59). Based on the gene location, protein size, and predicted membrane association, it is tempting to speculate that Xfas53 gp41, gp42, and gp43 are involved in superinfection exclusion despite the complete lack of primary structural similarity.

**Xfas53- and Xfp1-related myophage elements in five *X. fastidiosa* strains.** Prophages related to Xfas53 are present in each of the draft (Ann1 and Dixon) and complete (Temecula, M23, M12, and 9a5c) *X. fastidiosa* genome entries (Fig. 3B and 5). These include previously identified prophage elements Xfp1 and Xfp2 (in strain 9a5c); Xpd1 (in strain Temecula); Xop2, Xop3, and Xop9 (in strain Ann1); and Xap1 and Xap5 (in strain Dixon) (7, 8, 18, 53, 61). Related prophage elements were also identified in the unpublished but complete *X. fastidiosa* genome assemblies from strains M12 (located between nucleotides 1207697 and 1254095 of GenBank entry GI:167964044) and strain M23 (located between nucleotides 1297360 and 1344391 of GenBank entry GI:182630682). These phages can be divided into three groups based on alignments with each other: Xfas53 and Xop2; Xop9; and Xfp1, Xpf2, Xpd1, Xap1, Xap5, Xop3, M12, and M23 (Fig. 5). Xop2 and Xfas53 align over their entire length, indicating that Xop2 is expected to produce a similar, P22-related podophage. Xop2 is shorter than Xfas53, due primarily to deletions in the replication and regulation module, resulting in the loss of several genes, and therefore may not be functional. The most significant difference in gene content between Xop2 and Xfas53 is in the lysis cassette. The Xop2 lysis cassette has an atypical gene order in that the endolysin (XfasoDRAFT_4319) is encoded by the gene preceding the holin (XfasoDRAFT_4320), which is followed by the genes for Rz (XfasoDRAFT_4321) and Rz1 (not annotated). Similarly arranged lysis cassettes are present in Xfp1 and related phages (see below). Pairwise nucleotide alignments between Xfas53 and the prophage elements reveal that they all have left arms related to that of Xfas53, from *gene01* (integrase) to *gene23* (primase) (Fig. 5). In Xop9, this includes coding regions from the integrase (XfasoDRAFT_3154) to the gp22 BroN domain protein (XfasoDRAFT_3171) (Fig. 5 and data not shown). The right half of Xop9 is not related to any of the other phages (Fig. 5) and has virion structural genes encoding proteins distantly related to those from *Actinobacillus* myophage Aaphi23 (reference 49 and data not shown).

**Xfp1-like prophage.** Eight of the prophage elements (Xfp1, Xfp2, Xpd1, Xop3, Xap1, Xap5, M12, and M23) were found to be over 90% identical to each other over their entire length (Fig. 3 and 5). M12 has a significant inverted region and may be defective. For ease of reference, this group of prophage elements will be referred to as "Xfp1-like" prophage elements. The similarity between Xfas53 and the Xfp1-like prophage elements is restricted to the left half of the genome, encompassing the gp01 integrase equivalent to gp23. Immediately following *gene23*, there is an ~800-bp noncoding region (Fig. 3). DNA alignments suggest that recombination occurred in the distal portion of *gene23* (data not shown).

The right half of the Xfp1-like prophages, encoding the lysis and virion morphogenesis proteins, is unrelated to Xfas53 and is organized like lambdoid prophages. However, the tail assembly genes are more closely related to P2, suggesting that the progenitor of these phages was a hybrid between a P22-like phage and a P2-like phage (7, 8). The tail assembly cluster of these prophages encodes homologues of P2 D, X, U, FII, FI, I, J, W, and V. Virions produced by these elements would be expected to demonstrate contractile-tailed (myophage) morphology. At 624 residues, the predicted major capsid protein of these phages (corresponding to PD1106 in Xpd1) is larger than

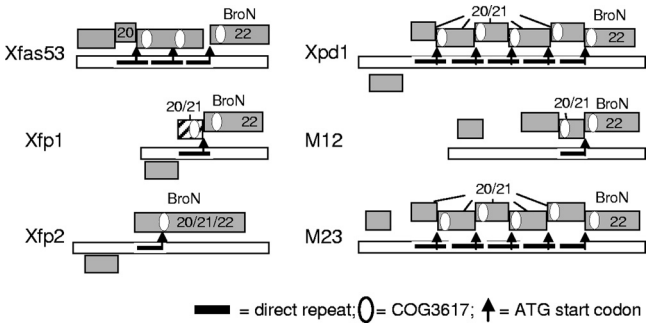


FIG. 6. Expansion of direct repeats encoding potential DNA binding proteins. Details of regions in Xfas53 and related prophages consisting of one, three, or five copies of a conserved 229- to 383-nucleotide region (indicated with a black bar). The unrelated gene immediately to the left of this region is also shown. In Xfas53, the coding region comprised by the repeats encodes gp20 and gp21. The repeat-encoded ATG start codon (indicated with an arrow) serves as the start codon of gp21 and gp22. The presence of COG3617 domains is indicated with ovals, and BroN (pfam02498) domains are also indicated. Genes not annotated in the original GenBank entries are hatched.

that from typical phages. The major capsid proteins from several phages, for example, *Stenotrophomonas* phage S1, are synthesized from similarly sized precursors (23). Like S1 and lactococcal bacteriophage c2, the Xpd1 major capsid protein contains a prohead protease domain, suggesting that the protein is self-processed and could also contain a scaffolding domain (15, 23). The Xfp1-like phages possess homologues of lambda *nu1*, *A*, and *B*, encoding the lambda TerS, TerL, and portal proteins, respectively. As in Xop2, the lysis cassette of the Xfp1-like prophage elements has an atypical gene arrangement, with the endolysin gene preceding the holin gene. The endolysin (XF0707 in Xfp1) is a glycoside hydrolase and is followed by XF0708, a homologue of the class I (3-TMD) holin from *Burkholderia* phage Bcep781 (56). The holin gene is followed by the *Rz* equivalent (XF0709 in Xfp1). Either the Rz1 equivalent in each of these prophages is not annotated (in Xfp1, Xfp2, and Xpd1) in the current version of the bacterial genome, or the incorrect start codon is annotated (in M12 and M23), resulting in the predicted protein being truncated from the amino terminus. Rz1 equivalent genes are frequently annotated incorrectly, due to their unusual gene arrangement (55). Interestingly, while the Rz equivalent of Xpd1 is unrelated to known Rz equivalents, the Rz1 equivalent is a homologue of PRD1 P37, which has been demonstrated to be an authentic Rz1 (35).

**Multiple BroN/COG3617 domain proteins and direct tandem repeats encoding multiple COG3617 domain proteins in Xfas53 and the Xfp1-like prophage elements.** The region of Xfas53 encoding gp20 and gp21 has an unusual organization, being composed of 2.7 copies of a 377-nucleotide direct tandem repeat (Fig. 6). These repeats are 87% identical. Variations of this repeat are present in each of the related *Xylella* prophages (Fig. 6). The prophages in M23 and Temecula each contain 4.7 copies of the repeat unit. In Xfp1, Xfp2, and M12, there is only a single, partial, and more degenerate copy of the repeat unit. The Xfas53 repeats are 87% identical with 5% of the differences attributed to mismatches and 7% of the differences attributed to insertions or deletions (indels). Each of the repeat units has an ATG served by a strong Shine-Dalgarno sequence predicted to serve in translational initiation. This ATG is located in the 3′ half of the repeat, such that the nucleotide sequence of the repeat is staggered in relation to the predicted coding region. The last repeat is truncated at this ATG, which serves as the start codon for the *gene22* equivalents. For Xfp2, the single-repeat-related sequence starts at this ATG and a fusion between the *gene21* and *gene22* equivalents is predicted to be expressed (XF2506). Five versions of the repeat-encoded protein are predicted to be encoded by Xpd1 and M23.

The repeat-encoded proteins are not related to proteins of known function, although most contain an amino-terminal conserved domain (COG3617) annotated as a prophage antirepressor (Fig. 6). Because of the indels, the amino acid sequences of the repeat-encoded proteins are less conserved than the nucleotide sequences. gp22 and the Xfp1-like prophage element equivalents each contain this same amino-terminal COG3617 conserved domain, which also encompasses a BroN (pfam02498) domain (Table 2). Additionally, gp8 and the Xfp1-like prophage element equivalents contain either two or three copies of the COG3617/BroN domains (Table 2). Not surprisingly, all of these proteins share some similarity to each other, due to the presence of the common COG3617 domain. Because the BroN domain defines a DNA binding domain found in numerous regulatory elements, it may be that these

TABLE 2. Xfas53 gp8, gp20/gp21, and gp22 homologues in *X. fastidiosa* prophage[a]

| Phage | gp8 gene product; number of conserved domains: identity of conserved domains | gp20/gp21 gene product(s); DNA repeat number, repeat length; protein conserved domains | gp22 gene product; conserved domains |
|---|---|---|---|
| Xfas53 | gp08; 2: BroN pfam02498/COG3617 | gp20 gp21; 2.7 copies, 381 bp; COG3617 | gp22; BroN pfam02498/COG3617 |
| Xfp2 | XF2524; 2: BroN pfam02498/COG3617 | XF2506[a]; one copy, 229 bp; BroN pfam02498/COG3617 | XF2506[a]; BroN, pfam02498/COG3617 |
| Xfp1 | XF0684; 3: BroN pfam02498/COG3617 | ORF not annotated; 1 copy, 382 bp; COG3617 | XF0704; BroN pfam02498/COG3617 |
| Xpd1 | PD1133; 3: BroN pfam02498/COG3617 | PD1120, PD1119, PD1118, PD1117; 4.7 copies, 383 bp; COG3617 | PD1116; BroN pfam02498/COG3617 |
| M12 | Xfasm12_1177; 2: BroN pfam02498/COG3617 | Xfasm12_1162; 1 copy, 269 bp; none | Xfasm12_1161; BroNpfam02498/COG3617 |
| M23 | Xfasm23_1205; 3: BroN pfam02498/COG3617 | Xfasm23_1186, Xfasm23_1185, Xfasm23_1184, Xfasm23_1183, Xfasm23_1182; 4.7 copies, 383 bp; COG3617 | Xfasm23_1180; BroN pfam02498/COG3617 |

[a] The Xfas52 gp20, gp21, and gp22 equivalents are present as a single open reading frame in Xfp2.

proteins are also regulatory elements that interact directly with DNA.

## DISCUSSION

Despite the extensive evidence of the involvement of phage in the genome diversity of *X. fastidiosa* and previous reports of the visualization of phage particles, this is the first report of the successful propagation of phage on *X. fastidiosa*. The extended time that it takes to prepare the plating culture (5 to 6 days) and the long incubation time required to visualize plaques (5 to 6 days) add to the challenge of manipulating *Xylella* phage. *Xylella* most commonly exists as an asymptomatic endophyte, showing disease symptoms on relatively few susceptible hosts (12). Xylem fluids are the most nutrient poor of all plant tissues, and *Xylella* bacteria are slow growing. Interestingly, the rate of adsorption of phage Xfas53 to strain Temecula was found to be extraordinarily low, with an experimentally determined adsorption rate constant of $6.8 \times 10^{-12}$ ml cell$^{-1}$ min$^{-1}$. This value is the lowest reported for a phage with its plaque-permissive host, 100- to 1,000-fold lower than those reported for other well-studied phages such as lambda PaPa (which lacks its side tail fibers) (24) at $1.3 \times 10^{-9}$ ml cell$^{-1}$ min$^{-1}$ and a reconstructed wild-type (wt) lambda phage at $9.9 \times 10^{-9}$ ml cell$^{-1}$ min$^{-1}$ (51). In the case of phage P22, a phage whose virion structural proteins are related to those of Xfas53, this constant has been reported as $3.4 \times 10^{-9}$ ml cell$^{-1}$ min$^{-1}$ with sensitive *Salmonella* hosts (39). The majority of models suggest that a high adsorption rate is a trait of improved fitness (1, 51). Given the unique ecology of *X. fastidiosa*, a xylem-limited plant pathogen, there may be benefits to slow adsorption. Growth is limited to individual xylem vessels unless connecting pit membranes are breached (16, 46). The transport characteristics of phage in xylem are unknown, but it is unlikely that phage virions can breach intact xylem pit membrane, considering that even small plant viruses require active transport through pit membranes (46). Pit membranes are constructed of plant cell wall material, and *X. fastidiosa* is able to traffic between vessels due to the production of plant cell wall-degrading enzymes (12). If *X. fastidiosa* phages require physiologically active host cells to facilitate movement through the xylem, slow adsorption might be advantageous in a situation where complete elimination of host bacteria in a vessel would result in reduced overall phage fecundity. Another effect of rapid phage adsorption in a xylem vessel would be an increased chance of progeny phage loss due to superinfection in the confined, host-limited space. Similar confined, host-limited space would also be confronted in the insect mouth parts where *Xylella* can reproduce. Recent experiments using double-layer agar plates as a model for reduced diffusion in biofilms suggest that while high adsorption rates are advantageous early in phage association with a biofilm, low adsorption rates are strongly selected for within the biofilm (22). It was proposed that low phage adsorption rates may increase overall phage fecundity in a biofilm-like environment by permitting longer periods of phage diffusion, increasing the effective size of the potential host pool, and reducing the probability of the phage adsorbing to an already infected cell or cell debris. There are conditions under which a reduced adsorption rate might be beneficial even to planktonic bacteria. Cyanophage AS-1, for example, exhibits light-dependent changes in adsorption kinetics (33). An interesting model was proposed suggesting that the lack of adsorption in the dark would result in phage remaining as free particles during the period when the host cells are not metabolically active, reducing the chances of the phage being grazed by nocturnal predators of bacteria (17).

A feature of the evolution of Xfas53 and related prophages is the expansion of BroN (pfam02498) and COG3617 domain-containing proteins. Bro (*b*aculovirus *r*epeat *o*pen reading frame) domain proteins are widespread in *Baculovirus* and other large DNA viruses, including phages (5). Bro domain proteins typically have at least two conserved domains with the amino-terminal domain, BroN, contributing the DNA-interacting domain and the specialized function being conferred by various C-terminal domains (29). Because the COG3617 regions include all of the BroN domains, it is tempting to speculate that COG3617 corresponds to a more distant lineage of BroN proteins also involved in DNA binding. BroN domain proteins often have C-terminal domains of various functions including endonucleases, protein-interacting domains, and additional DNA-interacting domains, as well as multiple conserved domains of unknown function (29). The evolution of Bro domain proteins is characterized by extensive lineage-specific domain expansion followed by domain shuffling. The tandem arrays that generate the multiple COG3617-containing proteins appear to be involved in expansion of COG3617 domain-containing proteins. The widespread occurrence of these proteins in *X. fastidiosa* prophage strongly suggests that they may play a pivotal and global role in *X. fastidiosa* phage biology. Increased expression of the Xfp1 Bro domain proteins was detected following heat shock treatment (18). We have conducted UV and mitomycin C induction experiments (using conditions which induce prophages of *Escherichia coli*, *Burkholderia* spp., and *Pseudomonas* spp.) with several *X. fastidiosa* lysogens, including the Xfas53 lysogen, without evidence of increased phage production or culture lysis (data not shown). Heat induction experiments with strain 53 did not show an increase in titer above that observed for the uninduced culture. The absence of lysogenic induction, despite the constitutive release of a basal level of virions in liquid culture, is puzzling. The Xfas53 cI repressor homologue appears to be a typical LexA family phage repressor, with conserved serine and lysine residues critical for protease activity. One possibility is that UV and mitomycin C treatments do not induce the SOS response required for prophage induction. Nevertheless, *Xylella* does encode homologues of *E. coli* RecA, LexA, and exonuclease V and components of the RecBCD complex known to participate in the repair of double-strand breaks produced during DNA replication (53, 61). Our understanding of prophage induction based on *E. coli* phage biology may not be applicable to *X. fastidiosa* phage systems.

Where does the obviously hybrid relationship of Xfas53 to the *Xylella* prophages fit into our current understanding of phage evolution? The capacity of phage to form hybrids has been known for some time. Hybrid phages are not particularly common, although the relationship between siphophage lambda and podophage P22 has long been recognized (38). The abundance of hybrid phages in *Xylella* implies that specific conditions are present that promote hybrid phage formation. *X. fastidiosa* strains carry multiple copies of closely related

prophages and multiple combinations of hybrid phages, a situation that can serve only to promote large-scale genomic rearrangements as well as the formation of new hybrid phage permutations (18). It was suggested that the larger number of genomic rearrangements of Xf-CVC and Xf-ALS than of Xf-PD and Xf-OLS may be the result of the presence of duplicated prophage elements in the former two strains, thus providing indirect evidence of the impact of prophage in reorganization of the *X. fastidiosa* genome (18). The Xfp1-like prophages themselves are of an ancient hybrid origin, having head assembly proteins related to lambda and tail assembly proteins related to those from P2 (8, 18). The extremely high DNA identity between Xfas53 and these myophage elements, however, indicates that this hybrid is of a very recent origin. One possibility is that these recombination events frequently generate nonfunctional prophages, which would also account for the lack of prophage induction. The isolation and propagation of Xfas53, a functional phage of recent hybrid origin, demonstrate that at least some of these combinations are in fact capable of forming viable virions.

## ACKNOWLEDGMENTS

We are grateful to Steven E. Lindow and Donald A. Cooksey, for providing *Xylella* strains. We thank Marie-Anne Van Sluys for providing us with the assembled Ann1 and Dixon prophage DNA sequences.

This work was supported by grants from the Texas Pierce's Disease Research and Education Program (to C.F.G.) and grant EF 0523951 from the National Science Foundation (to R.Y.) and funding from Texas AgriLife Research.

## REFERENCES

1. **Abedon, S. T., and R. R. Culler.** 2007. Bacteriophage evolution given spatial constraint. J. Theor. Biol. **248:**111–119.
2. **Berry, J., E. J. Summer, D. K. Struck, and R. Young.** 2008. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Mol. Microbiol. **70:**341–351.
3. **Bhattacharyya, A., S. Stilwagen, N. Ivanova, M. D'Souza, A. Bernal, A. Lykidis, V. Kapatral, I. Anderson, N. Larsen, T. Los, G. Reznik, E. Selkov, Jr., T. L. Walunas, H. Feil, W. S. Feil, A. Purcell, J. L. Lassez, T. L. Hawkins, R. Haselkorn, R. Overbeek, P. F. Predki, and N. C. Kyrpides.** 2002. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. Proc. Natl. Acad. Sci. U. S. A. **99:**12403–12408.
4. **Bhattacharyya, A., S. Stilwagen, G. Reznik, H. Feil, W. S. Feil, I. Anderson, A. Bernal, M. D'Souza, N. Ivanova, V. Kapatral, N. Larsen, T. Los, A. Lykidis, E. Selkov, Jr., T. L. Walunas, A. Purcell, R. A. Edwards, T. Hawkins, R. Haselkorn, R. Overbeek, N. C. Kyrpides, and P. F. Predki.** 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res. **12:**1556–1563.
5. **Bideshi, D. K., S. Renault, K. Stasiak, B. A. Federici, and Y. Bigot.** 2003. Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. J. Gen. Virol. **84:**2531–2544.
6. **Brodie, R., R. L. Roper, and C. Upton.** 2004. JDotter: a Java interface to multiple dotplots generated by dotter. Bioinformatics **20:**279–281.
7. **Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brussow.** 2003. Prophage genomics. Microbiol. Mol. Biol. Rev. **67:**238–276.
8. **Casjens, S.** 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. **49:**277–300.
9. **Casjens, S., and W. M. Huang.** 1982. Initiation of sequential packaging of bacteriophage P22 DNA. J. Mol. Biol. **157:**287–298.
10. **Casjens, S., D. A. Winn-Stapley, E. B. Gilcrease, R. Morona, C. Kuhlewein, J. E. Chua, P. A. Manning, W. Inwood, and A. J. Clark.** 2004. The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. J. Mol. Biol. **339:**379–394.
11. **Casjens, S. R., and E. B. Gilcrease.** 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol. Biol. **502:**91–111.
12. **Chatterjee, S., R. P. Almeida, and S. Lindow.** 2008. Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. Annu. Rev. Phytopathol. **46:**243–271.
13. **Chen, J., and E. L. Civerolo.** 2008. Morphological evidence for phages in *Xylella fastidiosa*. Virol. J. **5:**75.
14. **Chen, J., R. Groves, E. L. Civerolo, M. Viveros, M. Freeman, and Y. Zheng.** 2005. Two *Xylella fastidiosa* genotypes associated with almond leaf scorch disease on the same location in California. Phytopathology **95:**708–714.
15. **Cheng, H., N. Shen, J. Pei, and N. V. Grishin.** 2004. Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. Protein Sci. **13:**2260–2269.
16. **Choat, B., A. R. Cobb, and S. Jansen.** 2008. Structure and function of bordered pits: new discoveries and impacts on whole-plant hydraulic function. New Phytol. **177:**608–625.
17. **Clokie, M. R., and N. H. Mann.** 2006. Marine cyanophages and light. Environ. Microbiol. **8:**2074–2082.
18. **de Mello Varani, A., R. C. Souza, H. I. Nakaya, W. C. de Lima, L. G. Paula de Almeida, E. W. Kitajima, J. Chen, E. Civerolo, A. T. Vasconcelos, and M. A. Van Sluys.** 2008. Origins of the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PLoS One **3:**e4059.
19. **Doddapaneni, H., J. Yao, H. Lin, M. A. Walker, and E. L. Civerolo.** 2006. Analysis of the genome-wide variations among multiple strains of the plant pathogenic bacterium *Xylella fastidiosa*. BMC Genomics **7:**225.
20. **Efrony, R., I. Atad, and E. Rosenberg.** 2009. Phage therapy of coral white plague disease: properties of phage BA3. Curr. Microbiol. **58:**139–145.
21. **Farinha, M. A., and A. M. Kropinski.** 1997. Overexpression, purification, and analysis of the c1 repressor protein of *Pseudomonas aeruginosa* bacteriophage D3. Can. J. Microbiol. **43:**220–226.
22. **Gallet, R., Y. Shao, and I. N. Wang.** 2009. High adsorption rate is detrimental to bacteriophage fitness in a biofilm-like environment. BMC Evol. Biol. **9:**241.
23. **Garcia, P., C. Monjardin, R. Martin, C. Madera, N. Soberon, E. Garcia, A. Meana, and J. E. Suarez.** 2008. Isolation of new *Stenotrophomonas* bacteriophages and genomic characterization of temperate phage S1. Appl. Environ. Microbiol. **74:**7552–7560.
24. **Hendrix, R. W., and R. L. Duda.** 1992. Bacteriophage lambda PaPa: not the mother of all lambda phages. Science **258:**1145–1148.
25. **Hendson, M., A. H. Purcell, D. Chen, C. Smart, M. Guilhabert, and B. Kirkpatrick.** 2001. Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. Appl. Environ. Microbiol. **67:**895–903.
26. **Hernandez-Martinez, R., H. S. Costa, C. K. Dumenyo, and D. A. Cooksey.** 2006. Differentiation of strains of *Xylella fastidiosa* infecting grape, almonds, and oleander using a multiprimer PCR assay. Plant Dis. **90:**1382–1388.
27. **Hill, B. L., and A. H. Purcell.** 1995. Multiplication and movement of *Xylella fastidiosa* within grapevine and four other plants. Phytopathology **85:**1368–1372.
28. **Hofer, B., M. Ruge, and B. Dreiseikelmann.** 1995. The superinfection exclusion gene (*sieA*) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. J. Bacteriol. **177:**3080–3086.
29. **Iyer, L. M., E. V. Koonin, and L. Aravind.** 2002. Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSES transcription factors. Genome Biol. **3:**RESEARCH0012.
30. **Iyer, L. M., E. V. Koonin, D. D. Leipe, and L. Aravind.** 2005. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. Nucleic Acids Res. **33:**3875–3896.
31. **Jackson, E. N., F. Laski, and C. Andres.** 1982. Bacteriophage P22 mutants that alter the specificity of DNA packaging. J. Mol. Biol. **154:**551–563.
32. **Johnson, J. E., and W. Chiu.** 2007. DNA packaging and delivery machines in tailed bacteriophages. Curr. Opin. Struct. Biol. **17:**237–243.
33. **Kao, C. C., S. Green, B. Stein, and S. S. Golden.** 2005. Diel infection of a cyanobacterium by a contractile bacteriophage. Appl. Environ. Microbiol. **71:**4276–4279.
34. **Kropinski, A. M.** 2000. Sequence of the genome of the temperate, serotype-converting, Pseudomonas aeruginosa bacteriophage D3. J. Bacteriol. **182:**6066–6074.
35. **Krupovic, M., V. Cvirkaite-Krupovic, and D. H. Bamford.** 2008. Identification and functional analysis of the *Rz*/*Rz1*-like accessory lysis genes in the membrane-containing bacteriophage PRD1. Mol. Microbiol. **68:**492–503.
36. **Kwan, T., J. Liu, M. Dubow, P. Gros, and J. Pelletier.** 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J. Bacteriol. **188:**1184–1187.
37. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
38. **Lavigne, R., D. Seto, P. Mahadevan, H. W. Ackermann, and A. M. Kropinski.** 2008. Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. Res. Microbiol. **159:**406–414.
39. **Lindberg, A. A., M. Sarvas, and P. H. Makela.** 1970. Bacteriophage attachment to the somatic antigen of *Salmonella*: effect of O-specific structures in leaky R mutants and S, T1 hybrids. Infect. Immun. **1:**88–97.
40. **Lobocka, M. B., D. J. Rose, G. Plunkett III, M. Rusin, A. Samojedny, H. Lehnherr, M. B. Yarmolinsky, and F. R. Blattner.** 2004. Genome of bacteriophage P1. J. Bacteriol. **186:**7032–7068.
41. **Lu, M. J., Y. D. Stierhof, and U. Henning.** 1993. Location and unusual membrane topology of the immunity protein of the *Escherichia coli* phage T4. J. Virol. **67:**4905–4913.

42. **Lukashin, A. V., and M. Borodovsky.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
43. **Maillou, J., and B. Dreiseikelmann.** 1990. The *sim* gene of *Escherichia coli* phage P1: nucleotide sequence and purification of the processed protein. Virology **175:**500–507.
44. **Monteiro-Vitorello, C. B., M. C. de Oliveira, M. M. Zerillo, A. M. Varani, E. Civerolo, and M. A. Van Sluys.** 2005. *Xylella* and *Xanthomonas* Mobil'omics. OMICS **9:**146–159.
45. **Olia, A. S., A. Bhardwaj, L. Joss, S. Casjens, and G. Cingolani.** 2007. Role of *gene* 10 protein in the hierarchical assembly of the bacteriophage P22 portal vertex structure. Biochemistry **46:**8776–8784.
46. **Opalka, N., C. Brugidou, C. Bonneau, M. Nicole, R. N. Beachy, M. Yeager, and C. Fauquet.** 1998. Movement of rice yellow mottle virus between xylem cells through pit membranes. Proc. Natl. Acad. Sci. U. S. A. **95:**3323–3328.
47. **Park, T., D. K. Struck, C. A. Dankenbring, and R. Young.** 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J. Bacteriol. **189:**9135–9139.
48. **Ranade, K., and A. R. Poteete.** 1993. Superinfection exclusion (*sieB*) genes of bacteriophages P22 and lambda. J. Bacteriol. **175:**4712–4718.
49. **Resch, G., E. M. Kulik, F. S. Dietrich, and J. Meyer.** 2004. Complete genomic nucleotide sequence of the temperate bacteriophage Aa Phi 23 of *Actinobacillus actinomycetemcomitans*. J. Bacteriol. **186:**5523–5528.
50. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
50a. **Schwartz, M.** 1975. Reversible interaction between coliphage lambda and its receptor protein. J. Mol. Biol. **99:**185–201.
51. **Shao, Y., and I. N. Wang.** 2008. Bacteriophage adsorption rate and optimal lysis time. Genetics **180:**471–482.
52. **Sherald, J. L., S. S. Hearon, S. J. Kostka, and D. L. Morgan.** 1983. Sycamore leaf scorch: culture and pathogenicity of fastidious xylem-limited bacteria from scorch-affected trees. Plant Dis. **67:**849–852.
53. **Simpson, A. J., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, et al.** 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. Nature **406:**151–159.
54. **Summer, E. J.** 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. Methods Mol. Biol. **502:**27–46.
55. **Summer, E. J., J. Berry, T. A. Tran, L. Niu, D. K. Struck, and R. Young.** 2007. *Rz*/*Rz*1 lysis gene equivalents in phages of Gram-negative hosts. J. Mol. Biol. **373:**1098–1112.
56. **Summer, E. J., C. F. Gonzalez, M. Bomer, T. Carlile, A. Embry, A. M. Kucherka, J. Lee, L. Mebane, W. C. Morrison, L. Mark, M. D. King, J. J. LiPuma, A. K. Vidaver, and R. Young.** 2006. Divergence and mosaicism among virulent soil phages of the *Burkholderia cepacia* complex. J. Bacteriol. **188:**255–268.
57. **Summer, E. J., C. F. Gonzalez, T. Carlisle, L. M. Mebane, A. M. Cass, C. G. Savva, J. LiPuma, and R. Young.** 2004. *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. J. Mol. Biol. **340:**49–65.
58. **Tran, T. A., D. K. Struck, and R. Young.** 2007. The T4 RI antiholin has an N-terminal signal anchor release domain that targets it for degradation by DegP. J. Bacteriol. **189:**7618–7625.
59. **Vander Byl, C., and A. M. Kropinski.** 2000. Sequence of the genome of *Salmonella* bacteriophage P22. J. Bacteriol. **182:**6472–6481.
60. **van der Wilk, F., A. M. Dullemans, M. Verbeek, and J. F. van den Heuvel.** 1999. Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrthosiphon pisum*. Virology **262:**104–113.
61. **Van Sluys, M. A., M. C. de Oliveira, C. B. Monteiro-Vitorello, C. Y. Miyaki, L. R. Furlan, L. E. Camargo, A. C. da Silva, D. H. Moon, M. A. Takita, E. G. Lemos, M. A. Machado, et al.** 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J. Bacteriol. **185:**1018–1026.
62. **Wang, I. N., D. L. Smith, and R. Young.** 2000. Holins: the protein clocks of bacteriophage infections. Annu. Rev. Microbiol. **54:**799–825.
63. **Wu, H., L. Sampson, R. Parr, and S. Casjens.** 2002. The DNA site utilized by bacteriophage P22 for initiation of DNA packaging. Mol. Microbiol. **45:**1631–1646.
64. **Xu, M., A. Arulandu, D. K. Struck, S. Swanson, J. C. Sacchettini, and R. Young.** 2005. Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science **307:**113–117.
65. **Xu, M., D. K. Struck, J. Deaton, I. N. Wang, and R. Young.** 2004. A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. Proc. Natl. Acad. Sci. U. S. A. **101:**6415–6420.